TECOLOTE RESEARCH, INC.

Bridging Engineering and Economics

Since 1973

Common Errors When Using Risk Simulation Tools

Alfred Smith, Dr. Shu-Ping Hu

June 2005

■ Los Angeles ■ Washington, D.C. ■ Boston ■ Chantilly ■ Huntsville ■ Dayton ■ Santa Barbara

■ Albuquerque ■ Colorado Springs ■ Ft. Meade ■ Ft. Monmouth ■ Montgomery ■ Ogden ■ Patuxent River ■ Pensacola ■ San Diego

■ Cleveland ■ Denver ■ New Orleans ■ Oklahoma City ■ Silver Spring ■ Warner Robins AFB ■ Vandenberg AFB

TECOLOTE RESEARCH, INC.

Objective

The objective of this paper is to:

- **Familiarize analysts with modeling considerations when developing point (or reference) estimates**
- **Alert analysts to several common mistakes when estimating risk distribution inputs (mean, standard deviation)**
- **Provide tips on how to make the best use of any of the most popular risk simulation tools (Crystal Ball, @Risk, ACE)**
- **Demonstrate that if the appropriate care is taken, you will obtain the "same" result from any of the tools**

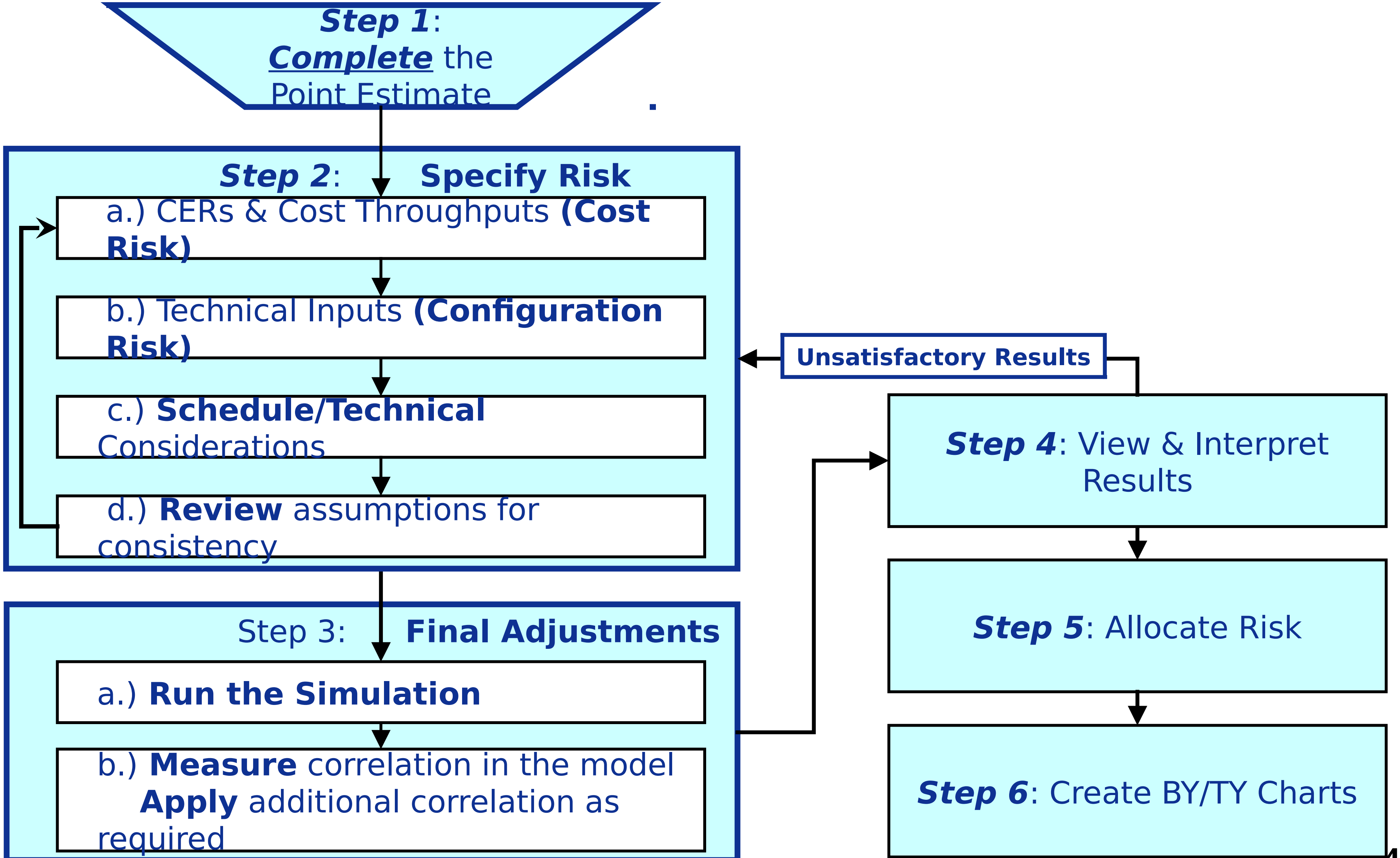
TECOLOTE RESEARCH, INC.

Outline

- Review cost risk analysis approach regardless of tool
- Common analytical errors
- Modeling Considerations
- Some detail on correlation
- All along the way, show examples
- Concluding Observations

TECOLOTE RESEARCH, INC.

Cost Risk Analysis Approach



Common Inflation Errors

- **Incorrect application of inflation is a common error in spreadsheets** (and a prime motivation for the creation of ACE)
  - Data is normalized to create a CER
  - CER inputs must therefore be normalized before used in the CER

WBS/CES Description	Unique ID	Eqn	FY	FY 2003
Space System NR				$480,484.07
Program Management/Systems Engineering	PMSE	1.487*(PLNR+SCNR)^0.841	1992	$78,844.45
Payload (P/L) Non Recurring	PLNR			$125,388.99
Payload IA&T				$18,766.74
Integration, Assembly, Test and Checkout (IA		850.764 + 0.159 * PLPME	1992	$17,959.81
Software Integration		.28*PLSW	2001	$806.93
Spacecraft (S/C) Non Recurring	SCNR			$276,250.63
Integration, assembly, Test and Checkout (IAT				$39,476.21
Spacecraft IA&T		850.764 + 0.159 * SCPME	1992	$38,654.01

- 1.487*(PLNR+SCNR)^0.841 **needs to be implemented as** 1.487*(PLNR**/Infl92to03**+SCNR**/Infl92to03**)^0.841***Infl92to03**
- 850.764 + 0.159 * SCPM **needs to be implemented as** 850.764***Infl92to03** + 0.159 * SCPM

FireSat Case Study

	Point Estimate	CER/Thruput	Risk
Ground Segment Operatioons	66,220		
Ground Seg Software (**SW**)	22,000	220 * KSLOC	$N(1, 0.25^2)$
Facilities	3,960	.18 * SW	$N(1, 0.25^2)$
Equipment	17,820	.81 * SW	$N(1, 0.25^2)$
Logistics	3,300	.15 * SW	$N(1, 0.25^2)$
Systems Level	19,140		
Management	3,960	.18 * SW	$N(1, 0.25^2)$
Systems engineering	6,600	.3 * SW	$N(1, 0.25^2)$
Product Assurance	3,300	.15 * SW	$N(1, 0.25^2)$
Integration and Test	5,280	.24 * SW	$N(1, 0.25^2)$
Ground Software **KSLOC**	100	100	T(0.95, 1, 2)

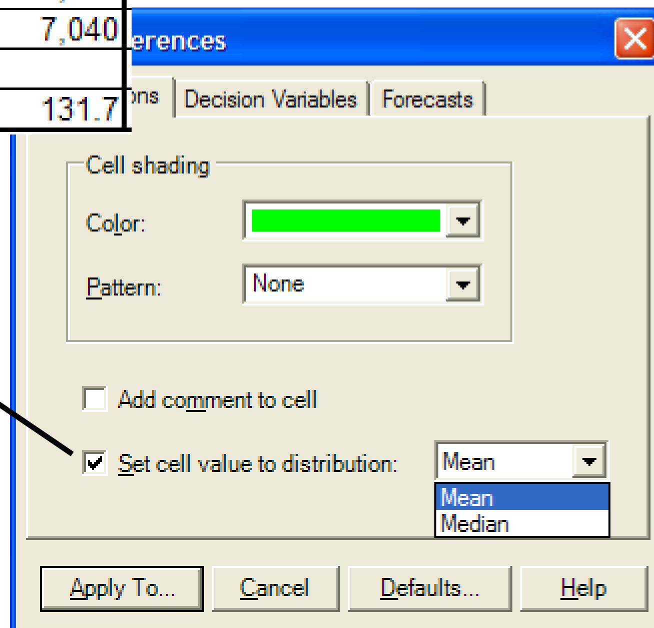
Correlation Matrix

	GSSW	FAC	Equip	Logis	Mgmt	SE	PA	Integ
GSSW	1.0	0.2	0.2	0.2	0.2	0.2	0.2	0.2
FAC		1.0	0.2	0.2	0.2	0.2	0.2	0.2
Equip			1.0	0.2	0.2	0.2	0.2	0.2
Logis				1.0	0.2	0.2	0.2	0.2
Mgmt					1.0	0.2	0.2	0.2
SE						1.0	0.2	0.2
PA							1.0	0.2
Integ								1.0

- **Software estimated as a function of KSLOC**
- **All other items estimated as a factor of Software**
- **Risk distributions assigned to KSLOC and all factors**
- **Additional Correlation “layered” on the model**

What is the "mean"

			Display		Simulation Result		
	Factor	Point Estimate	Forecast (CB Mean)	Analytic Mean	ACE Mean	CB Mean	@Risk Mean
Ground Segment Operatioons		$66,220	87,189.7	**87,917**	87,860	87,873	87,967
Ground Segment Software	220.00	$22,000	28,966.7	28,967	28,950	28,954	28,976
Facilities	0.18	$3,960	5,214.0	5,279	5,275	5,275	5,287
Equipment	0.81	$17,820	23,463.0	23,756	23,740	23,742	23,768
Logistics	0.15	$3,300	4,345.0	4,399	4,399	4,398	4,403
Systems Level		$19,140	25,201.0	**25,516**	25,497	25,504	25,532
Management	0.18	$3,960	5,214.0	5,279	5,278	5,275	5,283
Systems engineering	0.30	$6,600	8,690.0	8,799	8,778	8,796	8,807
Product Assurance	0.15	$3,300	4,345.0	4,399	4,396	4,399	4,402
Integration and Test	0.24	$5,280	6,952.0	7,039	7,045	7,035	7,040
Ground KSLOC (X)		100	131.7		131.7	131.7	131.7



- **CB cell preference for display can be misleading**
  - In this study, only 2 cells display the true mean
  - Do not assume display results represent the mean
- **$E(X*Y) = E(X)*E(Y) + \rho*\sigma_X*\sigma_Y$**
  For Facilities, E(Facilities) = E(0.18* 220*KSLOC*e_1* e_2)
  = 0.18*220*131.667*(1*1 + 0.2*0.25*0.25) = **5279.3**
- **All simulation tools match analytic mean within 0.1%**

What is the "Standard Deviation"

- **The standard deviation of a factor CER is often underestimated if a risk distribution is also specified for the input variable.**
- **For a factor CER: Y = a*X*e, the standard deviation of Y is given by**

$$\boxed{\begin{aligned} V(X\varepsilon) &= E(X\varepsilon)^2 - (E(X\varepsilon))^2 = E(X^2\varepsilon^2) - (E(X))^2(E(\varepsilon))^2 \\ &= E(X^2 - \mu_X^2)E(\varepsilon^2 - \mu_\varepsilon^2) + E(X^2)\mu_\varepsilon^2 - \mu_X^2\mu_\varepsilon^2 + \mu_X^2 E(\varepsilon^2) - \mu_X^2\mu_\varepsilon^2 \\ &= \sigma_X^2\sigma_\varepsilon^2 + (E(X^2) - \mu_X^2)\mu_\varepsilon^2 + \mu_X^2(E(\varepsilon^2) - \mu_\varepsilon^2) \\ &= \sigma_X^2\sigma_\varepsilon^2 + \sigma_X^2\mu_\varepsilon^2 + \mu_X^2\sigma_\varepsilon^2 \end{aligned}}$$

Stdev(SW) = Stdev(220*KSLOC*e)

$= (220)\{(\sigma_x^2)(\sigma_e^2) + (\sigma_x^2)(\mu_e^2) + (\mu_x^2)(\sigma_e^2)\}^{0.5}$ (KSLOC is denoted by X)

$= (220)\{(24.18)^2(0.25)^2 + (24.18)^2*(1)^2 + (131.67)^2(0.25)^2)\}^{0.5}$

= **9084** which is 25% greater than (220)(131.67) (0.25) = **7242**

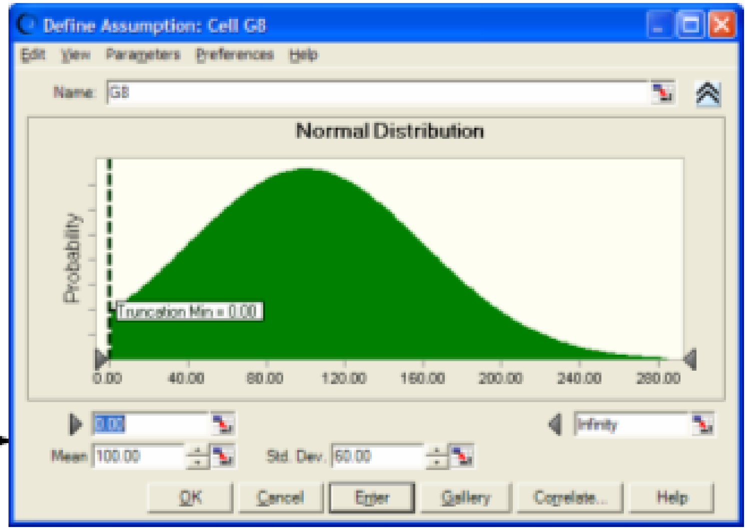
- **All simulation models estimate the SW standard deviation to be very close to the theoretical number,** i.e., 9,084

Common Problems Specifying Distributions

- **Many users prefer mean and stdev to define the risk distribution.**
  - CB & @Risk allows this, but problematic to remain consistent with changes
  - Must adjust all risk specifications if point estimate, units or FY changes
  - Methods available to allow risk distribution to scale with point estimate
- **Interpret CER result differently (mean, median, mode, other?)**
  - In general, CERs that are modeled with LN distribution produce the median
  - All other CERs assumed to produce the mode
- **"Schedule/Penalty" factors in ACE can make it problematic to establish similar distribution in other tools**
- **The Beta distribution is a special case**
  - requires four parameters (low, high, alpha, beta)
  - generally only have the point estimate and assumed upper/lower bounds
  - triangular distribution based with same low/mode/high defines the range & provides sufficient information to uniquely define the beta distribution.
  - DoD had this simplification built into ACE when RI$K was conceived.
  - CB and @Risk can not define beta distributions in this manner.

Truncation

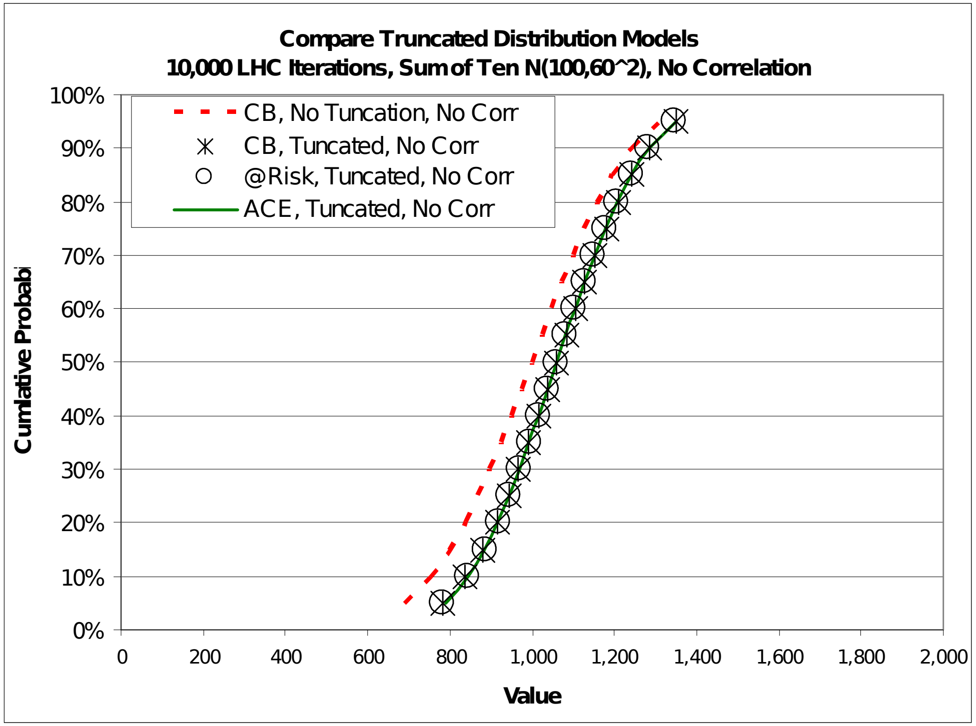
	Normal Distribution Assumptions	Truncated Normal Distribution Assumptions
Total	1000	1000
Element 1	100	100
Element 2	100	100
Element 3	100	100
Element 4	100	100
Element 5	100	100
Element 6	100	100
Element 7	100	100
Element 8	100	100
Element 9	100	100
Element 10	100	100

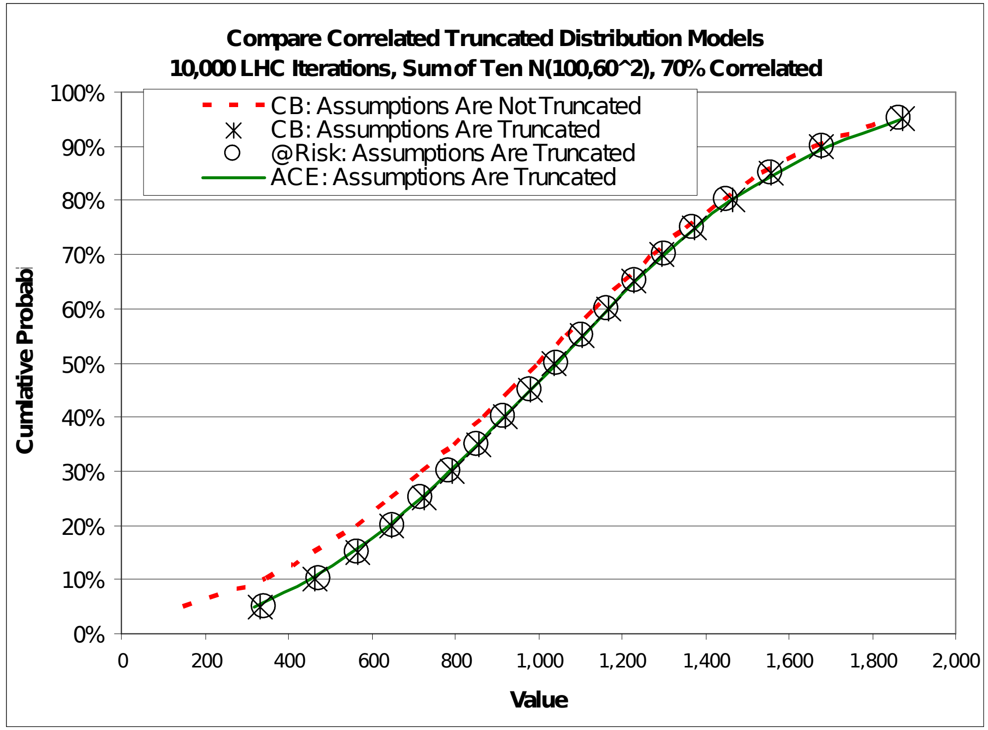


Number of Trials: 10,000
LHC Sample Size: 10,000
Initial Seed Value: 999

- **ACE truncates all assumptions at zero by default**
- **CB & @Risk can model truncation**
  - Be careful to set the assumption not just the graphical display
- **All tools preserve total number of iterations**
- **Does anyone think a CER result or technical input should be permitted to go below zero?**

Impact of Truncation Reduces as Correlation Increases



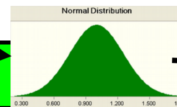
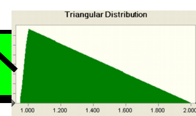


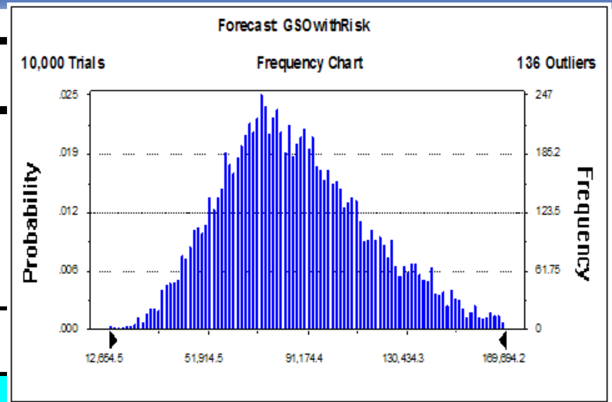
- **ACE cannot (by design) model non-truncated distributions**
- **Becomes less important as correlation increases, particularly 50% and above**

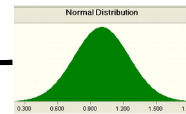
Learning

- **Cumulative Average Cost learning curve: T1* Qty ^(b+1) * ε**
- **Many users are tempted to put distributions on the T1, the slope and also the CER as a whole**
- **If the T1 and slope were derived together through regression analysis, then there should only be risk on the CER as a whole**
- **If the T1 and slope are independent guesses, then putting risk on both of them is appropriate but:**
  - One should not also apply risk to the CER as a whole
  - The T1 and slope risk should be highly, negatively correlated

Risk on Top Of Risk

	Factor	Risk Assumption	Display Forecast (CB Mean)
Ground Segment Operatioons			87,189.7
Ground Segment Software	220.00		28,966.7
Facilities	0.18	1.000	5,214.0
Equipment	0.81	1.000	23,463.0
Logistics	0.15	1.000	4,345.0
Systems Level			25,201.0
Management	0.18	1.000	5,214.0
Systems engineering	0.30	1.000	8,690.0
Product Assurance	0.15	1.000	4,345.0
Integration and Test	0.24	1.000	6,952.0
Ground KSLOC (X)			131.7





Common mistakes include:

- **CER risk forgotten**
- **Incorrect flow**
- **Failure to link to forecast cells**

Other Modeling Issues (1/2)

- **User-Defined Functions**
  - Excel can be somewhat arbitrary in the execution order of UDFs and CB can deliver unexpected results.
  - Reference 2 discusses this issue in detail and proposes VBA methods to control the sequence of execution
- **Log-Error CERs**
  - A log-linear CER (i.e., a * X^b) generates the median rather than mean.
  - If the mean is required, the estimate must be adjusted (see "Ping Factor" Reference 3, 4, and 5).
- **Correlation Matrices**
  - Reference 6 states "Although not widely documented, it appears that after 60 – 80 correlations have been entered the (CB) behavior can be erratic."
  - We have no knowledge of this problem in @Risk. ACE does not have this problem.

Other Modeling Issues (2/2)

- **The conventions used by ACE for specifying distributions and correlation are confusing to some**
- **CB may not alert the user when the point estimate is outside the defined risk distribution bounds**
- **Duplicating an entire Excel worksheet that contains “local” range names can cause simulation tools to become confused**
- **Multiple models within the same Excel workbook can lead to confusing reports**
- **CB documentation alerts user that 250 is upper limit for correlations (see reference 6 for clever work around)**
- **Inconsistent global settings (random seed, sampling method, iterations, partitions) across models will influence results**
- **ACE session file containing all the models (90 rows in the spreadsheet) was 112k. CB/@Risk models/reports spanned 2 workbooks, over 20 spreadsheets consuming 25Meg**

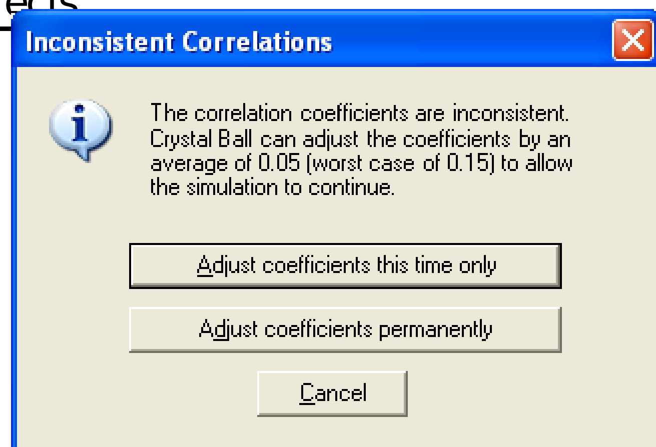
Project Cost Risk Analysis Example (Ref 7 David T. Hulett)

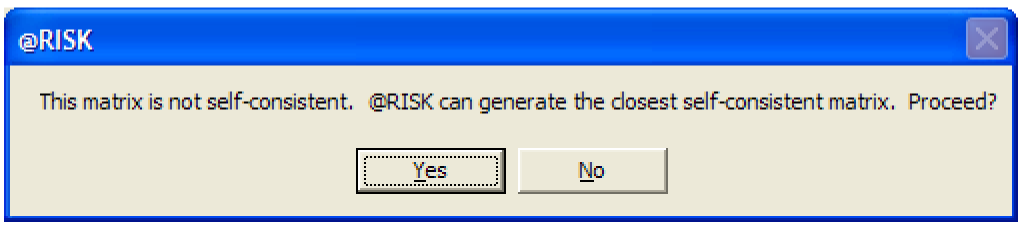
Simple estimate to illustrate the impact of applying an inconsistent correlation matrix in CB

This example can be found on the CB web site at http://www.crystalball.com/articles/hulett.html)

WBS	Point Estimate	Triangular Bounds	
Total	$25,200	Low	High
Design	$1,500	$1,200	$2,000
Equipment	$5,000	$4,500	$6,000
Foundation, Structure	$7,000	$6,200	$8,600
Piping, Elect., HVAC	$1,000	$800	$1,600
Labor	$4,700	$4,500	$5,900
Indirects	$6,000	$5,000	$6,700

Entered Correlations						
	Des	Equip	Struc	HVAC	Labor	Ind
Design	1.00	0.80	0.70	0.90	0.80	0.90
Equip		1.00	0.50	0.40	0.60	0.50
Structure			1.00	0.40	0.70	0.80
HVAC				1.00	0.20	0.40
Labor					1.00	0.60
Indirects						1.00





TECOLOTE RESEARCH, INC.

Correlation

- **Need to know the correlation assumptions *used* by the tool rather than those that are *entered*.**
- **CB and @Risk do different things when the entered correlation matrix is incomplete or "inconsistent".**

Entered Correlations

	Des	Equip	Struc	HVAC	Labor	Ind
Design	1.00	0.80	0.70	0.90	0.80	0.90
Equip		1.00	0.50	0.40	0.60	0.50
Structure			1.00	0.40	0.70	0.80
HVAC				1.00	0.20	0.40
Labor					1.00	0.60
Indirects						1.00

Correlations Used By CB

	Des	Equip	Struc	HVAC	Labor	Ind
Design	1.00	0.72	0.69	0.75	0.69	0.78
Equip		1.00	0.47	0.40	0.58	0.49
Structure			1.00	0.37	0.65	0.75
HVAC				1.00	0.22	0.42
Labor					1.00	0.60
Indirects						1.00

>10% less than entered
>5% less than entered
>5% more than entered

Entered Correlations

	Des	Equip	Struc	HVAC	Labor	Ind
Design	1.00	0.80	0.70	0.90	0.80	0.90
Equip		1.00	0.50	0.40	0.60	0.50
Structure			1.00	0.40	0.70	0.80
HVAC				1.00	0.20	0.40
Labor					1.00	0.60
Indirects						1.00

Correlations Used by @Risk

	Des	Equip	Struc	HVAC	Labor	Ind
Design	1.00	0.67	0.58	0.75	0.67	0.75
Equip		1.00	0.42	0.33	0.50	0.42
Structure			1.00	0.33	0.58	0.67
HVAC				1.00	0.17	0.33
Labor					1.00	0.50
Indirects						1.00

TECOLOTE RESEARCH, INC.

Applying the same Correlation in Different Tools

Used CB ADJUSTED Correlation Inputs (not the entered ones)

Use ACE Correlation Wizard to enter Correlation

Choosing the Vector for ACE

	Des	Equip	Struc	HVAC	Labor	Ind
Design	1.000	0.723	0.693	0.755	0.693	0.780
Equip	0.723	1.000	0.470	0.399	0.583	0.489
Structure	0.693	0.470	1.000	0.366	0.654	0.748
HVAC	0.755	0.399	0.366	1.000	0.223	0.417
Labor	0.693	0.583	0.654	0.223	1.000	0.596
Indirects	0.780	0.489	0.748	0.417	0.596	1.000
Average	0.774	0.611	0.655	0.527	0.625	0.672

	WBS/CES	Stren	85	86	87	88	89	90
85	Design	D	1.000	0.723	0.693	0.755	0.693	0.780
86	Equipment	0.723		1.000	0.501	0.546	0.501	0.564
87	Foundation,	0.693			1.000	0.523	0.480	0.541
88	Piping, Elect.,	0.755				1.000	0.523	0.589
89	Labor	0.693					1.000	0.541
90	Indirects	0.780						1.000

Measured Simulated Correlations

	Des	Equip	Struc	HVAC	Labor	Ind
Design	1.00	0.73	0.71	0.79	0.71	0.76
Equip		1.00	0.48	0.40	0.59	0.49
Structure			1.00	0.38	0.67	0.76
HVAC				1.00	0.24	0.43
Labor					1.00	0.60
Indirects						1.00

WBS/CES	Row 85: Design	Row 86: Equipm	Row 87: Founda tion	Row 88: Piping, Elect.,	Row 89: Labor	Row 90: Indirec
Design	1.000	0.717	0.688	0.749	0.678	0.773
Equipment		1.000	0.500	0.536	0.490	0.553
Foundation, Structure			1.000	0.516	0.464	0.538
Piping, Elect., HVAC				1.000	0.508	0.584
Labor					1.000	0.525
Indirects						1.000

- **ACE estimates the correlation matrix from a single vector**
- **Select the vector by completing the matrix and find highest average**
- **ACE provides a correlation report. For CB and @Risk, extract iteration data to a sheet and use RANK and CORREL to calculate simulated correlations**

Basis for ACE Group Strength Approach to Correlation

The following is the proof that the cross-correlation between two Group elements can be approximated by the product of their user-defined Group strength values.

Consider a correlation group is composed of *n* elements, and they are specified with the group strengths ρ_1, ρ_2.... ρ_n The pairwise correlated (normal) items by the Cholesky's decomposition method are given by:

$$X_1 = \rho_1 * Z + \sqrt{1 - \rho_1^2} * Z_1$$

$$X_2 = \rho_2 * Z + \sqrt{1 - \rho_2^2} * Z_2$$

$$\ldots$$

$$X_n = \rho_n * Z + \sqrt{1 - \rho_n^2} * Z_n$$

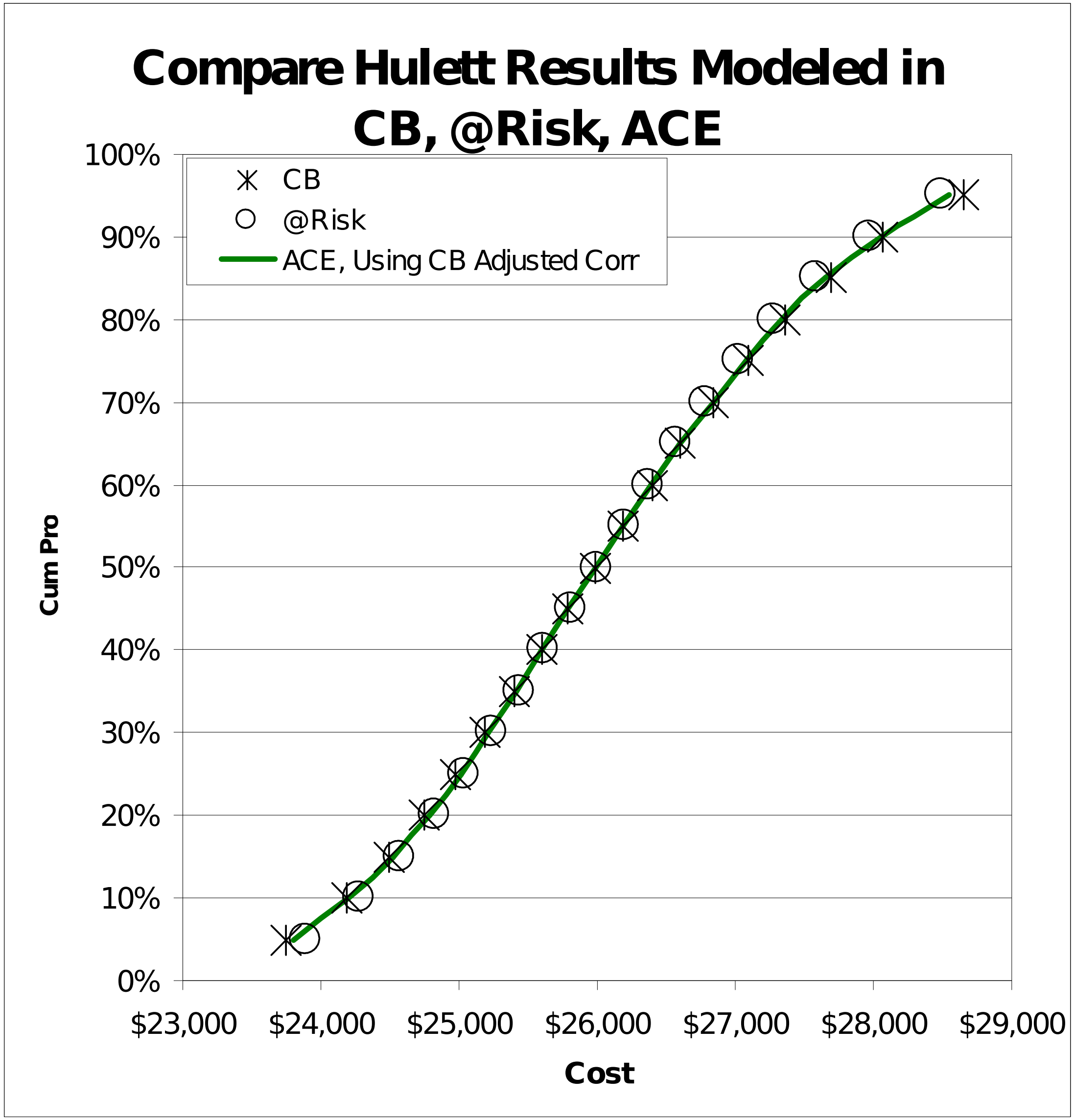
The correlation between X_i and X_j is then given by:

$$\rho(X_i, X_j) = \frac{Cov(X_i, X_j)}{\sigma_{x_i}\,\sigma_{x_j}} = \frac{Cov(\rho_i Z + \sqrt{1-\rho_i^2}\, Z_i,\ \rho_j Z + \sqrt{1-\rho_j^2}\, Z_j)}{\sigma_z\,\sigma_z}$$

$$= \frac{Cov(\rho_i Z, \rho_j Z)}{\sigma_z^2} = \frac{\rho_i\,\rho_j\,Cov(Z, Z)}{\sigma_z^2} = \rho_i\,\rho_j$$

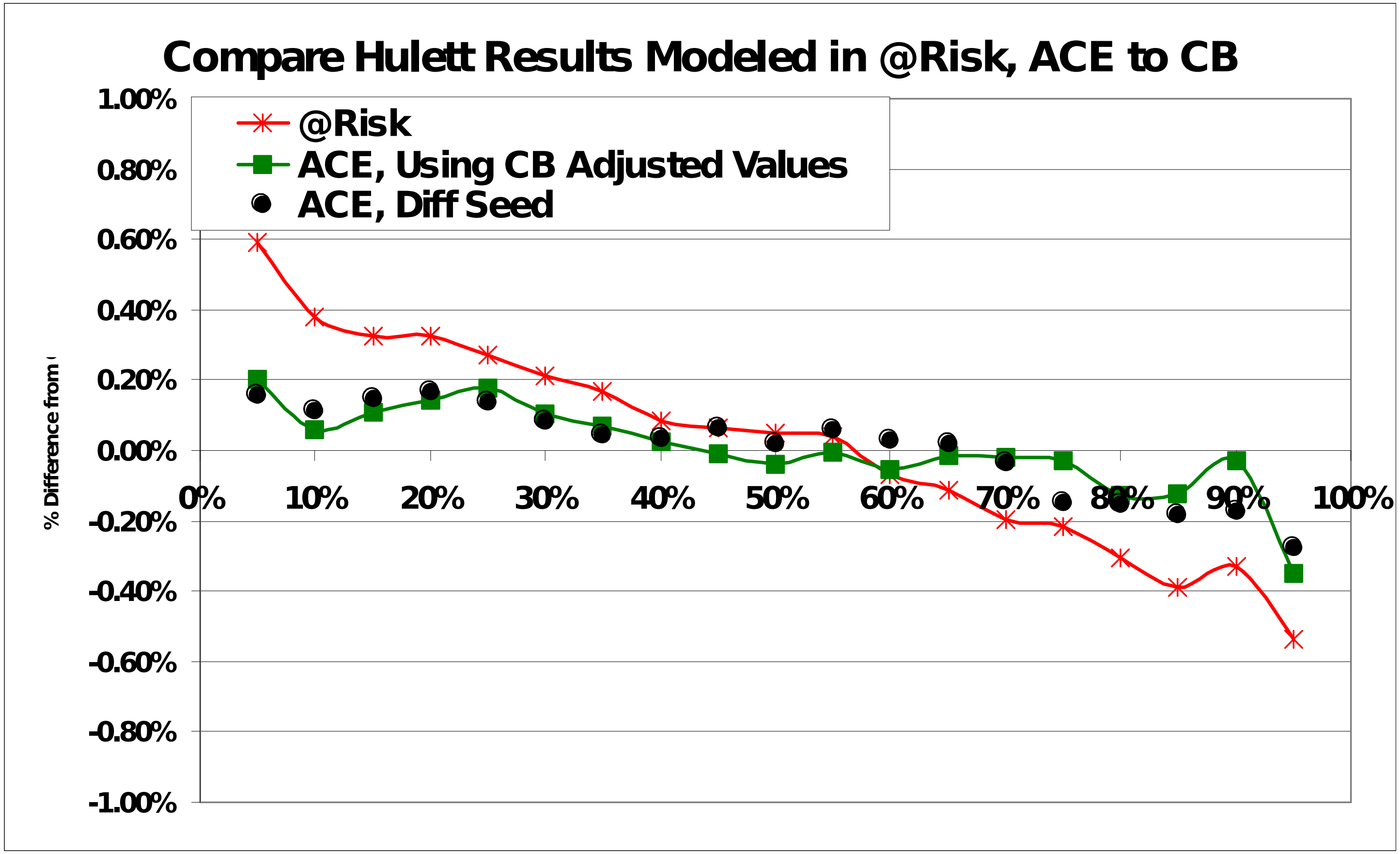
TECOLOTE RESEARCH, INC.

Compare CB, @Risk and ACE When Correlation Applied



- **For ACE, used the correlations as adjusted by CB**
- **@Risk made greater adjustments to the correlations**
- **Curiously, for this example, ACE results tend to be the average of CB and @Risk**
- **If correlations are removed, all three tools match almost identically**

Impact of Different Tool Adjustments To Inconsistent Correlation Matrix



- Lines illustrate difference between the tool and CB result
- ACE within 0.2%
- @Risk somewhat different due to different adjustment technique (back-up slides identify how each tool does it)

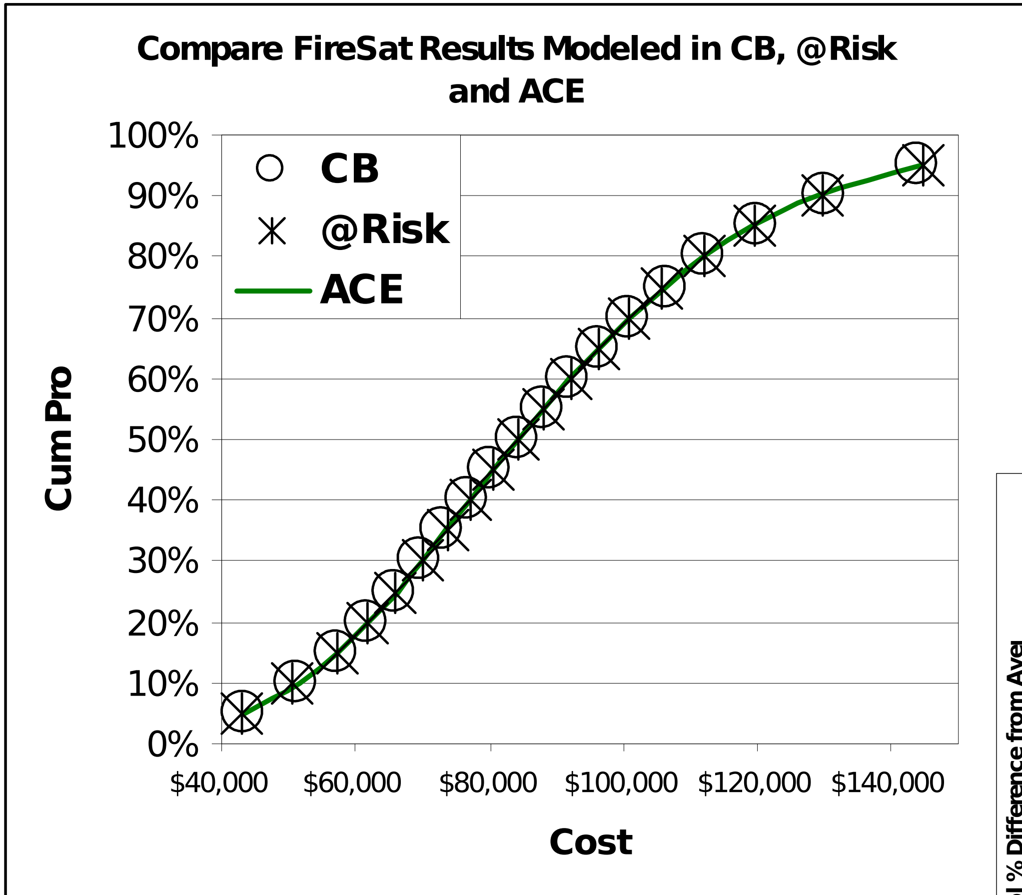
Layering Correlation on Top of Functional Correlation

	WBS/CES	Row 32: Ground Segment Software (SW)	Row 33: Facilities	Row 34: Equipment	Row 35: Logistics	Row 37: Management	Row 38: Systems engineering	Row 39: Product Assurance	Row 40: Integration and Test
32	Ground Segment Software (SW)	1.0000	0.7666	0.7653	0.7613	0.7649	0.7679	0.7676	0.7654
33	Facilities		1.0000	0.5867	0.5833	0.5870	0.5825	0.5915	0.5809
34	Equipment			1.0000	0.5888	0.5827	0.5852	0.5895	0.5829
35	Logistics				1.0000	0.5688	0.5829	0.5870	0.5806
37	Management					1.0000	0.5915	0.5825	0.5870
38	Systems engineering						1.0000	0.5940	0.5934
39	Product Assurance							1.0000	0.5888
40	Integration and Test								1.0000

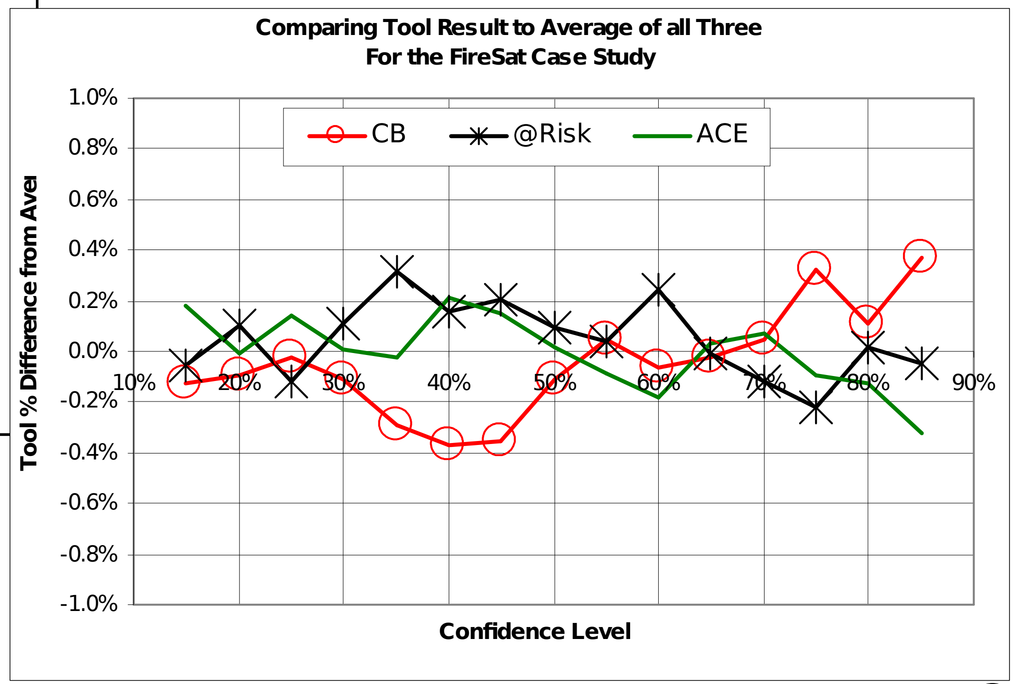
	WBS/CES	Row 32: Ground Segment Software (SW)	Row 33: Facilities	Row 34: Equipment	Row 35: Logistics	Row 37: Management	Row 38: Systems engineering	Row 39: Product Assurance	Row 40: Integration and Test
32	Ground Segment Software (SW)	1.0000	0.8059	0.8054	0.8014	0.8054	0.8073	0.8074	0.8057
33	Facilities		1.0000	0.7139	0.7077	0.7127	0.7085	0.7159	0.7081
34	Equipment			1.0000	0.7157	0.7121	0.7133	0.7194	0.7127
35	Logistics				1.0000	0.7008	0.7088	0.7139	0.7095
37	Management					1.0000	0.7170	0.7141	0.7146
38	Systems engineering						1.0000	0.7208	0.7193
39	Product Assurance							1.0000	0.7166
40	Integration and Test								1.0000

- **Recall that FireSat estimate, every row is a factor of SW**
  - Upper graphic illustrates functional correlation present before any added
- **Lower graphic illustrates the impact of “adding” 20% more correlation**
  - Potential to over specify correlation

All Tools Handle Functional plus Additional Correlation



- Simple 20 percent "layered" across all elements
- If risk flow, distributions, and correlations applied consistently, all tools match



Conclusions

- **Common errors when analyzing risk for multiplicative CERs:**
  - common to underestimate the mean and standard deviations for multiplicative CERs when risk distributions are also specified for the independent variables. (i.e. 43% in the Firesat case study!)
- **Excessive Correlation**
  - Can lead to errors in calculating the mean and stdev
  - Recommend measuring correlation present before adding more
  - tempting to defeat a cost model's functional correlation (one truth) and replace them with "observed correlations" at higher levels (another truth?)…creative methods available…. but is it right?
- **Common pitfalls when using analytic methods**
  - it is easy to get "lost in the math" even for simple estimates, especially when functional correlation is involved,
  - assuming the central limit theorem applies at parent WBS (i.e. totals tend to be normal distributions) is often incorrect
  - After much effort, often find it impossible to derive close-form solutions

Errors to Look For if Simulation Tools Fail to Match

- improper cost input normalization to non-linear CERs
- inappropriate analytic methods to estimate a combined (multiplicative error) CER and its input risk (rather than model each element separately)
- inconsistent risk distribution specification on CERS and their inputs
- different truncation assumptions
- inconsistent “risk flow” modeling in Excel based tools
- ignoring or defeating functional correlation
- failure to account for how the tool “adjusts” correlation
- failure to follow manufacturer’s guidance when applying correlation
- confusion over which cell in the Excel spreadsheet contains the “right” answer

2

TECOLOTE RESEARCH, INC.
Bridging Engineering and Economics
Since 1973

Backup Slides

@Risk User Guide Notes on Correlation

- **In general, a matrix is valid only if it is *positive semi-definite*. A positive semi-definite matrix has eigen values which are all greater than or equal to zero, and at least one eigenvalue that is greater than zero.**
- **If @RISK determines you have an invalid matrix, it will give you the option of letting @RISK generate the closest valid matrix to the entered invalid one. @RISK follows these steps to modify a matrix:**
- **1) Finds the smallest eigenvalue (E0)**
- 2) **"Shifts" the eigenvalues so that the smallest eigenvalue equals zero by adding the product of -E0 and the identity matrix (**I**) to the correlation matrix (**C**):** C' = C **- E0**I.
- **3) Divides the new matrix by 1 - E0 so that the diagonal terms equal:** C'' = **(1/1-E0)**C'

CB User Guide Notes on Correlation

- There is a practical limit of about 250 fully correlated assumptions (assumptions that are correlated to every other assumption) and up to about 1,000 serially correlated assumptions (assumptions that are correlated to one or two other assumptions) for each workbook.
- If you have a large number of correlated assumptions:
  - For coefficients that are close to zero and are nearly independent, remove the correlation.
  - For coefficients that are close to one, replace one of the assumptions with a formula in your spreadsheet.
- If Crystal Ball detects inconsistently correlated assumptions when a simulation is running, it first determines whether small adjustments to the correlation coefficients are possible. This process might take a long time, depending on the number of correlated assumptions. Crystal Ball displays the message “Examining the Correlation Coefficients”. **If you get this message, you should probably stop and redefine your correlations.**
- If small adjustments to the correlation coefficients are possible, a dialog appears, letting you decide whether to cancel the simulation or continue with the adjusted coefficients.

(no description on how CB adjusts an inconsistent matrix)

ACEIT Is Structured to Automate the Estimating Environment